



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Breece, Tim  
Hayenga, Kirk  
Rinderknecht, Ernst  
Vandlen, Richard  
Yansura, Daniel

(ii) TITLE OF INVENTION: PROCESS FOR PRODUCING RELAXIN

(iii) NUMBER OF SEQUENCES: 40

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Mr. Walter H. Dreger  
(B) STREET: 4 Embarcadero Center, Suite 3400  
(C) CITY: San Francisco  
(D) STATE: California  
(E) COUNTRY: USA  
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/080,354  
(B) FILING DATE: 21-JUN-1993  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Dreger, Walter H.  
(B) REGISTRATION NUMBER: 24,190  
(C) REFERENCE/DOCKET NUMBER: A-58117/WH

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (415) 781-1989  
(B) TELEFAX: (415) 398-3249

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asp Ser Trp Met Glu Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val  
1 5 10 15

Arg Ala Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser  
20 25

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Lys Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

a'  
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr  
1 5 10 15

Lys Arg Ser Leu Ala Arg Phe Cys  
20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Lys Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Arg Arg Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Lys Lys Arg  
1 5 10

a'  
cont.

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Arg Arg Lys  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 231 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 1..231

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG AAA AAG AAT ATC GCA TTT CTT CTT AAA CGG GAC TCA TGG ATG GAG 48  
 Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg Asp Ser Trp Met Glu  
 1 5 10 15

GAA GTT ATT AAA TTA TGC GGC CGC GAA TTA GTT CGC GCG CAG ATT GCC 96  
 Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala  
 20 25 30

ATT TGC GGC ATG AGC ACC TGG AGC AAA AGG AAA CCC ACT GGT TAT GGT 144  
 Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr Gly Tyr Gly  
 35 40 45

TCT CGA AAA AAG AGA CAA CTC TAC AGT GCA TTG GCT AAT AAA TGT TGC 192  
 Ser Arg Lys Lys Arg Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys  
 50 55 60

CAT GTT GGT TGT ACC AAA AGA TCT CTT GCT AGA TTT TGC 231  
 His Val Gly Cys Thr Lys Arg Ser Leu Ala Arg Phe Cys  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 77 amino acids

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Arg Asp Ser Trp Met Glu  
1 5 10 15  
Glu Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala  
20 25 30  
Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr Gly Tyr Gly  
35 40 45  
Ser Arg Lys Lys Arg Gln Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys  
50 55 60  
His Val Gly Cys Thr Lys Arg Ser Leu Ala Arg Phe Cys  
65 70 75

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 593 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 431..586

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC ATGAAAAATC TCATTGCTGA 60  
GTTGTTATTT AAGCTTGCCC AAAAAGAAGA AGAGTCGAAA GAACTGTGTG CGCAGGTAGA 120  
AGCTTTGGAG ATTATCGTCA CTGCAATGCT TCGCAATATG GCGCAAAATG ACCAACAGCG 180  
GTTGATTGAT CAGGTAGAGG GGGCGCTGTA CGAGGTAAAG CCCGATGCCA GCATTCTCTGA 240  
CGACGATACG GAGCTGCTGC GCGATTACGT AAAGAAGTTA TTGAAGCATC CTCGTCAGTA 300  
AAAAGTTAAT CTTTTCACAA GCTGTCATAA AGTTGTCACG GCCGAGACTT ATAGTCGCTT 360  
TGTTTTTATT TTTTAATGTA TTTGTACGCA AGTTCACGTA AAAAGGGTAT CTAGAGGTTG 420  
AGGTGATTTT ATG AAA AAG AAT ATC GCA TTT CTT CTT GCA TCT ATG TTC 469  
Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe  
1 5 10

GTT TTT TCT ATT GCT ACA AAT GCC TAT GCA GAC TCA TGG ATG GAG GAA 517  
Val Phe Ser Ile Ala Thr Asn Ala Tyr Ala Asp Ser Trp Met Glu Glu  
15 20 25

GTT ATT AAA TTA TGC GGC CGC GAA TTG GTA CGC GCG CAA ATA GCG ATA 565  
Val Ile Lys Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile  
30 35 40 45

TGC GGT ATG AGT ACA TGG AGT TGAAGAA 593  
Cys Gly Met Ser Thr Trp Ser  
50

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
1 5 10 15  
Ile Ala Thr Asn Ala Tyr Ala Asp Ser Trp Met Glu Glu Val Ile Lys  
20 25 30  
Leu Cys Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met  
35 40 45  
Ser Thr Trp Ser  
50

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1500 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 438..1238

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC ATGAAAAATC TCATTGCTGA 60

a' cont.

a  
cont.

AAG GAC TTA ATT TAT ACA CTT TAT TAT TGG AAA TCT TCA AGT TCA GGA 998  
 Lys Asp Leu Ile Tyr Thr Leu Tyr Tyr Trp Lys Ser Ser Ser Ser Gly  
 175 180 185  
 AAG AAA ACA GCC AAA ACA AAC ACT AAT GAG TTT TTG ATT GAT GTG GAT 1046  
 Lys Lys Thr Ala Lys Thr Asn Thr Asn Glu Phe Leu Ile Asp Val Asp  
 190 195 200  
 AAA GGA GAA AAC TAC TGT TTC AGT GTT CAA GCA GTG ATT CCC TCC CGA 1094  
 Lys Gly Glu Asn Tyr Cys Phe Ser Val Gln Ala Val Ile Pro Ser Arg  
 205 210 215  
 ACA GTT AAC CGG AAG AGT ACA GAC AGC CCG GTA GAG TGT ATG GGC CAG 1142  
 Thr Val Asn Arg Lys Ser Thr Asp Ser Pro Val Glu Cys Met Gly Gln  
 220 225 230 235  
 GAG AAA GGC CAA TTC AGA GAA ATA TTC TAC ATC ATT GGA GCT GTG GTA 1190  
 Glu Lys Gly Gln Phe Arg Glu Ile Phe Tyr Ile Ile Gly Ala Val Val  
 240 245 250  
 TTT GTG GTC ATC ATC CTT GTC ATC ATC CTG GCT ATA TCT CTA CAC TAAAATTCTC 1245  
 Phe Val Val Ile Ile Leu Val Ile Ile Leu Ala Ile Ser Leu His  
 255 260 265  
 ATGTTTGACA GCTTATCATC GATAAGCTTT AATGCGGTAG TTTATCACAG TTAAATTGCT 1305  
 AACGCAGTCA GGCACCGTGT ATGAAATCTA ACAATGCGGT CATCGTCATC CTCGGCACCG 1365  
 TCACCCTGGA TGCTGTAGGC ATAGGCTTGG TTATGCGGT ACTGCCGGGC CTCTTGCGGG 1425  
 ATATCGTCCA TTCCGACAGC ATGCCAGTC ACTATGGCGT GCTCCTAGCG CTATATGCGT 1485  
 TGATGCAATT TCTAT 1500

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser  
 1 5 10 15  
 Ile Ala Thr Asn Ala Tyr Ala Ser Gly Thr Thr Asn Thr Val Ala Ala  
 20 25 30  
 Tyr Asn Leu Thr Trp Lys Ser Thr Asn Phe Lys Thr Ile Leu Glu Trp  
 35 40 45  
 Glu Pro Lys Pro Val Asn Gln Val Tyr Thr Val Gln Ile Ser Thr Lys



50 55 60

Ser Gly Asp Trp Lys Ser Lys Cys Phe Tyr Thr Thr Asp Thr Glu Cys  
65 70 75 80

Asp Leu Thr Asp Glu Ile Val Lys Asp Val Lys Gln Thr Tyr Leu Ala  
85 90 95

Arg Val Phe Ser Tyr Pro Ala Gly Asn Val Glu Ser Thr Gly Ser Ala  
100 105 110

Gly Glu Pro Leu Tyr Glu Asn Ser Pro Glu Phe Thr Pro Tyr Leu Glu  
115 120 125

Thr Asn Leu Gly Gln Pro Thr Ile Gln Ser Phe Glu Gln Val Gly Thr  
130 135 140

Lys Val Asn Val Thr Val Glu Asp Glu Arg Thr Leu Val Arg Arg Asn  
145 150 155 160

Asn Thr Phe Leu Ser Leu Arg Asp Val Phe Gly Lys Asp Leu Ile Tyr  
165 170 175

Thr Leu Tyr Tyr Trp Lys Ser Ser Ser Ser Gly Lys Lys Thr Ala Lys  
180 185 190

Thr Asn Thr Asn Glu Phe Leu Ile Asp Val Asn Lys Gly Glu Asn Tyr  
195 200 205

Cys Phe Ser Val Gln Ala Val Ile Pro Ser Arg Thr Val Asn Arg Lys  
210 215 220

Ser Thr Asp Ser Pro Val Glu Cys Met Gly Gln Glu Lys Gly Gln Phe  
225 230 235 240

Arg Glu Ile Phe Tyr Ile Ile Gly Ala Val Val Phe Val Val Ile Ile  
245 250 255

Leu Val Ile Ile Leu Ala Ile Ser Leu His  
260 265

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Lys Lys Asn Ile Ala Phe Leu Leu Arg Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Lys Lys Asn Ile Ala Phe Leu Leu Arg Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Lys Lys Asn Ile Ala Phe Leu Leu Lys Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..42
- (D) OTHER INFORMATION: /note= "Complementary  
double-stranded binding between bases 5 and 42 to  
SEQ ID NO:16. "

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTAGAATTAT GAAAAAGAAT ATCGCATTTTCTTAAACG GG

42

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 4..41
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 4 and 41 to SEQ ID NO:15."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AGTCCCGTTT AAGAAGAAAT GCGATATTCT TTTTCATAAT T

41

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..42
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 5 and 42 to SEQ ID NO:18."

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 9..41

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTAGAATT ATG TTC CCA GCT ATG CCT CTA TCT AGT AAA CCG G  
Met Phe Pro Ala Met Pro Leu Ser Ser Lys Arg  
1 5 10

42

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

a!  
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Phe Pro Ala Met Pro Leu Ser Ser Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 4..41
- (D) OTHER INFORMATION: /note= "Complementary  
double-stranded binding between bases 4 and 41 to  
SEQ ID NO:17."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AGTCCCGTTT ACTAGATAGA GGCATAGCTG GGAACATAAT T

41

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 5..64

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..64
- (D) OTHER INFORMATION: /note= "Complementary  
double-stranded binding between bases 5 and 64 to  
SEQ ID NO:21."

a!  
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGCG CAG ATT GCC ATT TGC GGC ATG AGC ACC TGG AGC AAA AGG AAA CCC 49  
 Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro  
 1 5 10 15  
 ACT GGT TAT GGT TCT 64  
 Thr Gly Tyr Gly Ser  
 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Gln Ile Ala Ile Cys Gly Met Ser Thr Trp Ser Lys Arg Lys Pro Thr  
 1 5 10 15  
 Gly Tyr Gly Ser  
 20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 3..62
- (D) OTHER INFORMATION: /note= "Complementary  
 double-stranded binding between bases 3 and 62 to  
 SEQ ID NO:20."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGAGAACCAT AACCAGTGGG TTTCCTTTTG CTCCAGGTGC TCATGCCGCA AATGGCAATC 60  
 TG 62

a'  
cont.

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..50
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 5 and 50 with SEQ ID NO:24."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGCCACTCTG TCGGGTGCTG AACTGGTGA CGCTCTGCAG TTTGTTTGCG

50

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 6..50
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 6 and 50 with SEQ ID NO:23."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTCACCGCAA ACAAACTGCA GAGCGTCAAC CAGTTCAGCA CCGCAAGAGT

50

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

a' cont.

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..55
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 1 and 55 with SEQ ID NO:26."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGTCCCGAAA CTCTGTGCGG TGCTGAACTG GTTGACGCTC TGCAGTTTGT TTGCG 55

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 6..64
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 6 and 64 with SEQ ID NO:25."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTCACCGCAA ACAAAGTGA GAGCGTCAAC CAGTTCAGCA CCGCACAGAG TTTCGGGACC 60

TGCA 64

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..84
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 5 and 84 with SEQ ID NO:28."

a'  
ant.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTAGAATTAT GATGATTACT CTGCGCAAAC TTCCTCTGGC GGTGCGCGTC GCAGCGGGCG 60  
TAATGTCTGC TCAGGCCATG GCCA 84

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 5..84
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 5 and 84 with SEQ ID NO:27."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GATCTGGCCA TGGCCTGAGC AGACATTACG CCCGCTGCGA CGGCAACCGC CAGAGGAAGT 60  
TTGCGCAGAG TAATCATCAT AATT 84

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..52
- (D) OTHER INFORMATION: /note= "Complementary double-stranded binding between bases 1 and 52 with SEQ ID NO:30."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CAACTCTACA GTGCATTGGC TAATAAATGT TGCCATGTTG GTTGACCAA AA 52

(2) INFORMATION FOR SEQ ID NO:30:

a1  
cont.



- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION: 5..56  
(D) OTHER INFORMATION: /note= "Complementary  
double-stranded binding between bases 5 and 56  
with SEQ ID NO:29."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GATCTTTTGG TACAACCAAC ATGGCAACAT TTATTAGCCA ATGCACTGTA GAGTTGTGCA 60

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Arg Arg Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Asp Lys Lys Arg Thr Gly Tyr Gly Ser Arg Lys Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Lys Arg Lys Pro Thr Gly Tyr Gly Ser Arg Arg Arg Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 915 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..452

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GC GGC CGC GAA TTA GTT CGC GCG CAG ATT GCC ATT TGC GGC ATG AGC 47  
Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met Ser  
1 5 10 15

ACC TGG AGC AAA AGG TCT CTG AGC CAG GAA GAT GCT CCT CAG ACA CCT 95  
Thr Trp Ser Lys Arg Ser Leu Ser Gln Glu Asp Ala Pro Gln Thr Pro  
20 25 30

AGA CCA GTG GCA GAA ATT GTG CCA TCC TTC ATC AAC AAA GAT ACA GAA 143  
Arg Pro Val Ala Glu Ile Val Pro Ser Phe Ile Asn Lys Asp Thr Glu  
35 40 45

ACC ATA AAT ATG ATG TCA GAA TTT GTT GCT AAT TTG CCA CAG GAG CTG 191  
Thr Ile Asn Met Met Ser Glu Phe Val Ala Asn Leu Pro Gln Glu Leu  
50 55 60

AAG TTA ACC CTG TCT GAG ATG CAG CCA GCA TTA CCA CAG CTA CAA CAA 239  
Lys Leu Thr Leu Ser Glu Met Gln Pro Ala Leu Pro Gln Leu Gln Gln  
65 70 75

CAT GTA CCT GTA TTA AAA GAT TCC AGT CTT CTC TTT GAA GAA TTT AAG 287

His Val Pro Val Leu Lys Asp Ser Ser Leu Leu Phe Glu Glu Phe Lys  
80 85 90 95

AAA CTT ATT CGC AAT AGA CAA AGT GAA GCC GCA GAC AGC AGT CCT TCA 335  
Lys Leu Ile Arg Asn Arg Gln Ser Glu Ala Ala Asp Ser Ser Pro Ser  
100 105 110

GAA TTA AAA TAC TTA GGC TTG GAT ACT CAT TCT CGA AAA AAG AGA CAA 383  
Glu Leu Lys Tyr Leu Gly Leu Asp Thr His Ser Arg Lys Lys Arg Gln  
115 120 125

CTC TAC AGT GCA TTG GCT AAT AAA TGT TGC CAT GTT GGT TGT ACC AAA 431  
Leu Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr Lys  
130 135 140

AGA TCT CTT GCT AGA TTT TGC TGAGATGAAG CTAATTGTGC ACATCTCGTA 482  
Arg Ser Leu Ala Arg Phe Cys  
145 150

TAATATTAC ACATATTCTT AATGACATTT CACTGATGCT TCTATCAGGT CAATTCTCAT 542

GTTTGACAGC TTATCATCGA TAAGCTTTAA TGCGGTAGTT TATCACAGTT AAATTGCTAA 602

CGCAGTCAGG CACCGTGTAT GAAATCTAAG AATGCGCTCA TCGTCATCCT CGGCACCGTC 662

ACCCTGGATG CTGTAGGCAT AGGCTTGGTT ATCCCGGTAC TGCCGGGCCT CTTGCGGGAT 722

ATCGTCCATT CCGACAGCAT CGCCAGTCAC TATGCGGTGC TGCTAGCGCT ATATGCGTTG 782

ATGCAATTTT TATGCGCACC CGTTCTCGGA GCACTGTCCG ACCGCTTTGG CCGCCGCCCA 842

GTCCTGCTCG CTTGCTACT TGGAGCCACT ATCGACTACG CGATCATGGC GACCACACCC 902

GTCCTGTGGA TCC 915

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 150 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Arg Glu Leu Val Arg Ala Gln Ile Ala Ile Cys Gly Met Ser Thr  
1 5 10 15

Trp Ser Lys Arg Ser Leu Ser Gln Glu Asp Ala Pro Gln Thr Pro Arg  
20 25 30

Pro Val Ala Glu Ile Val Pro Ser Phe Ile Asn Lys Asp Thr Glu Thr  
35 40 45

Ile Asn Met Met Ser Glu Phe Val Ala Asn Leu Pro Gln Glu Leu Lys  
50 55 60

Leu Thr Leu Ser Glu Met Gln Pro Ala Leu Pro Gln Leu Gln Gln His  
65 70 75 80

Val Pro Val Leu Lys Asp Ser Ser Leu Leu Phe Glu Glu Phe Lys Lys  
85 90 95

Leu Ile Arg Asn Arg Gln Ser Glu Ala Ala Asp Ser Ser Pro Ser Glu  
100 105 110

Leu Lys Tyr Leu Gly Leu Asp Thr His Ser Arg Lys Lys Arg Gln Leu  
115 120 125

Tyr Ser Ala Leu Ala Asn Lys Cys Cys His Val Gly Cys Thr Lys Arg  
130 135 140

Ser Leu Ala Arg Phe Cys  
145 150

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 301 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 7..297

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AAGCTT ATG AAA TCT AAC AAT GCG CTC ATC GTC ATC CTC GGC ACC GTC 48  
Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val  
1 5 10

ACC CTG GAT GCT GTA GGC ATA GGC TTG GTT ATG CCG GTA CTG CCG GGC 96  
Thr Leu Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly  
15 20 25 30

CTC TTG CGG GAT ATC GTC CAT TCC GAC AGC ATC GCC AGT CAC TAT GGC 144  
Leu Leu Arg Asp Ile Val His Ser Asp Ser Ile Ala Ser His Tyr Gly  
35 40 45

GTG CTG CTA GCG CTA TAT GCG TTG ATG CAA TTT CTA TGC GCA CCC GTT 192  
Val Leu Leu Ala Leu Tyr Ala Leu Met Gln Phe Leu Cys Ala Pro Val  
50 55 60

CTC GGA GCA CTG TCC GAC CGC TTT GGC CGC CGC CCA GTC CTG CTC GCT 240

a'  
cont.

Leu Gly Ala Leu Ser Asp Arg Phe Gly Arg Arg Pro Val Leu Leu Ala  
65 70 75

TCG CTA CTT GGA GCG ACT ATC GAC TAC GCG ATC ATG GCG ACC ACA CCC  
Ser Leu Leu Gly Ala Thr Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro  
80 85 90

288

GTC CTG TGG ATCC  
Val Leu Trp  
95

301

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Lys Ser Asn Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu  
1 5 10 15  
Asp Ala Val Gly Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu  
20 25 30  
Arg Asp Ile Val His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu  
35 40 45  
Leu Ala Leu Tyr Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly  
50 55 60  
Ala Leu Ser Asp Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu  
65 70 75 80  
Leu Gly Ala Thr Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu  
85 90 95  
Trp

a  
concl.